

SEQUENCE LISTING

(1) GENERAL INFORMATION:

(i) APPLICANT: MELKI, JUDITH
MUNNICH, ARNOLD

(ii) TITLE OF INVENTION: SURVIVAL MOTOR NEURON (SMN) GENE: A GENE
FOR SPINAL MUSCULAR ATROPHY

(iii) NUMBER OF SEQUENCES: 57

(iv) CORRESPONDENCE ADDRESS:

- (A) ADDRESSEE: BIRCH, STEWART, KOLASCH AND BIRCH, LLP
(B) STREET: PO BOX 747
(C) CITY: FALLS CHURCH
(D) STATE: VA
(E) COUNTRY: USA
(F) ZIP: 22040-0747

(v) COMPUTER READABLE FORM:

- (A) MEDIUM TYPE: Floppy disk
(B) COMPUTER: IBM PC compatible
(C) OPERATING SYSTEM: PC-DOS/MS-DOS
(D) SOFTWARE: PatentIn Release #1.0, Version #1.30

(vi) CURRENT APPLICATION DATA:

- (A) APPLICATION NUMBER:
(B) FILING DATE:
(C) CLASSIFICATION:

(viii) ATTORNEY/AGENT INFORMATION:

- (A) NAME: FARACI, C. J.
(B) REGISTRATION NUMBER: 32,350
(C) REFERENCE/DOCKET NUMBER: 2121-110P

(ix) TELECOMMUNICATION INFORMATION:

- (A) TELEPHONE: (703) 205-8000
(B) TELEFAX: (703) 205-8050

(2) INFORMATION FOR SEQ ID NO:1:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 347 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:1:

AATTTTAA	TTTTTGTAG	AGACAGGGTC	TCATTATGTT	CCCCAGGGTG	GTGTCAAGCT	60
CCAGGTCTCA	AGTGATCCCC	CTACCTCCGC	CTCCCAAAGT	TGTGGGATTG	TAGGCATGAG	120
CCACTGCAAG	AAAACCTTAA	CTGCAGCCTA	ATAATTGTTT	TCTTTGGGAT	AACTTTTAA	180
GTACATTAAA	AGACTATCAA	CTTAATTTCT	GATCATATTT	TGTTGAATAA	AATAAGTAAA	240
ATGTCTTGTG	AACAAAATGC	TTTTTAACAT	CCATATAAAG	CTATCTATAT	ATAGCTATCT	300
ATGTCTATAT	AGCTATTTTT	TTTAACTTCC	TTTTATTTTC	CTTACAG		347

(2) INFORMATION FOR SEQ ID NO:2:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 444 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:2:

GTAGTCTGC	CAGCATTATG	AAAGTGAATC	TTACTTTTGT	AAAACTTTAT	GGTTTGTGGA	60
AAACAAATGT	TTTTGAACAG	TTAAAAAGTT	CAGATGTTAA	AAAGTTGAAA	GGTTAATGTA	120
AAACAATCAA	TATTAAAGAA	TTTTGATGCC	AAAAC TATTA	GATAAAAGGT	TAATCTACAT	180
CCCTACTAGA	ATTCTCATAC	TTAACTGGTT	GGTTATGTGG	AAGAAACATA	CTTTCACAAT	240
AAAGAGCTTT	AGGATATGAT	GCCATTTTAT	ATCACTAGTA	GGCAGACCAG	CAGACTTTTT	300
TTTATTGTGA	TATGGGATAA	CCTAGGCATA	CTGCACTGTA	CACTCTGACA	TATGAAGTGC	360
TCTAGTCAAG	TTTAACTGGT	GTCCACAGAG	GACATGGTTT	AACTGGAATT	CGTCAAGCCT	420
CTGGTTCTAA	TTTCTCATTT	GCAG				444

(2) INFORMATION FOR SEQ ID NO:3:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 347 base pairs
- (B) TYPE: nucleic acid

(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:3:

AATTTTAAA	TTTTTGTAG	AGACAGGGTC	TCATTATGTT	GCCCAGGGTG	GTGTCAAGCT	60
CCAGGTCTCA	AGTGATCCCC	CTACCTCCGC	CTCCCAAAGT	TGTGGGATTG	TAGGCATGAG	120
CCACTGCAAG	AAAACCTTAA	CTGCAGCCTA	ATAATTGTTT	TCTTTGGGAT	AACTTTTAAA	180
GTACATTAAA	AGACTATCAA	CTTAATTTCT	GATCATATTT	TGTTGAATAA	AATAAGTAAA	240
ATGCTTGTG	AACAAAATGC	TTTTTAACAT	CCATATAAAG	CTATCTATAT	ATAGCTATCT	300
ATATCTATAT	AGCTATTTTT	TTTAACTTCC	TTTTATTTTC	CTTACAG		347

(2) INFORMATION FOR SEQ ID NO:4:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 444 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:4:

GTAAGTCTGC	CAGCATTATG	AAAGTGAATC	TTACTTTTGT	AAAACCTTAT	GGTTTGTGGA	60
AAACAAATGT	TTTTGAACAG	TTAAAAAGTT	CAGATGTTAG	AAAGTTGAAA	GGTTAATGTA	120
AAACAATCAA	TATTAAAGAA	TTTTGATGCC	AAAACCTATTA	GATAAAAGGT	TAATCTACAT	180
CCCTACTAGA	ATTCTCATAC	TTAACTGGTT	GGTGTGTGG	AAGAAACATA	CTTTCACAAT	240
AAAGAGCTTT	AGGATATGAT	GCCATTTTAT	ATCACTAGTA	GGCAGACCAG	CAGACTTTTT	300
TTTATTGTGA	TATGGGATAA	CCTAGGCATA	CTGCACTGTA	CACTCTGACA	TATGAAGTGC	360
TCTAGTCAAG	TTTAACTGGT	GTCCACAGAG	GACATGGTTT	AACTGGAATT	CGTCAAGCCT	420
CTGGTTCTAA	TTTCTCATTT	GCAG				444

(2) INFORMATION FOR SEQ ID NO:5:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 25 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

- (A) DESCRIPTION: /desc = "SYNTHETIC DNA"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:5:

AGACTATCAA CTTAATTTCT GATCA

25

(2) INFORMATION FOR SEQ ID NO:6:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 24 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

- (A) DESCRIPTION: /desc = "SYNTHETIC DNA"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:6:

TAAGGAATGT GAGCACCTTC CTTC

24

(2) INFORMATION FOR SEQ ID NO:7:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 23 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

- (A) DESCRIPTION: /desc = "SYNTHETIC DNA"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:7:

GTAATAACCA AATGCAATGT GAA

23

(2) INFORMATION FOR SEQ ID NO:8:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 20 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

- (A) DESCRIPTION: /desc = "SYNTHETIC DNA"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:8:

CTAGAACACC CTTCTCACAG

20

(2) INFORMATION FOR SEQ ID NO:9:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 294 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:9:

Met	Ala	Met	Ser	Ser	Gly	Gly	Ser	Gly	Gly	Gly	Val	Pro	Glu	Gln	Glu	1	5	10	15
Asp	Ser	Val	Leu	Phe	Arg	Arg	Gly	Thr	Gly	Gln	Ser	Asp	Asp	Ser	Asp	20	25	30	
Ile	Trp	Asp	Asp	Thr	Ala	Leu	Ile	Lys	Ala	Tyr	Asp	Lys	Ala	Val	Ala	35	40	45	
Ser	Phe	Lys	His	Ala	Leu	Lys	Asn	Gly	Asp	Ile	Cys	Glu	Thr	Ser	Gly	50	55	60	
Lys	Pro	Lys	Thr	Thr	Pro	Lys	Arg	Lys	Pro	Ala	Lys	Lys	Asn	Lys	Ser	65	70	75	80
Gln	Lys	Lys	Asn	Thr	Ala	Ala	Ser	Leu	Gln	Gln	Trp	Lys	Val	Gly	Asp	85	90	95	

Lys Cys Ser Ala Ile Trp Ser Glu Asp Gly Cys Ile Tyr Pro Ala Thr
 100 105 110
 Ile Ala Ser Ile Asp Phe Lys Arg Glu Thr Cys Val Val Val Tyr Thr
 115 120 125
 Gly Tyr Gly Asn Arg Glu Glu Gln Asn Leu Ser Asp Leu Leu Ser Pro
 130 135 140
 Ile Cys Glu Val Ala Asn Asn Ile Glu Gln Asn Ala Gln Glu Asn Glu
 145 150 155 160
 Asn Glu Ser Gln Val Ser Thr Asp Glu Ser Glu Asn Ser Arg Ser Pro
 165 170 175
 Gly Asn Lys Ser Asp Asn Ile Lys Pro Lys Ser Ala Pro Trp Asn Ser
 180 185 190
 Phe Leu Pro Pro Pro Pro Pro Met Pro Gly Pro Arg Leu Gly Pro Gly
 195 200 205
 Lys Pro Gly Leu Lys Phe Asn Gly Pro Pro Pro Pro Pro Pro Pro Pro
 210 215 220
 Pro Pro His Leu Leu Ser Cys Trp Leu Pro Pro Phe Pro Ser Gly Pro
 225 230 235 240
 Pro Ile Ile Pro Pro Pro Pro Pro Ile Cys Pro Asp Ser Leu Asp Asp
 245 250 255
 Ala Asp Ala Leu Gly Ser Met Leu Ile Ser Trp Tyr Met Ser Gly Tyr
 260 265 270
 His Thr Gly Tyr Tyr Met Gly Phe Arg Gln Asn Gln Lys Glu Gly Arg
 275 280 285
 Cys Ser His Ser Leu Asn
 290

(2) INFORMATION FOR SEQ ID NO:10:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1582 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:10:

CGGGGCCCCA CGCTGCGCAC CCGCGGGTTT GCTATGGCGA TGAGCAGCGG CGGCAGTGGT	60
GGCGGCGTCC CGGAGCAGGA GGATTCCGTG CTGTTCCGGC GCGGCACAGG CCAGAGCGAT	120
GATTCTGACA TTTGGGATGA TACAGCACTG ATAAAAGCAT ATGATAAAGC TGTGGCTTCA	180
TTTAAGCATG CTCTAAAGAA TGGTGACATT TGTGAAACTT CGGGTAAACC AAAAACCACA	240
CCTAAAAGAA AACCTGCTAA GAAGAATAAA AGCCAAAAGA AGAATACTGC AGCTTCCTTA	300
CAACAGTGGA AAGTTGGGGA CAAATGTTCT GCCATTTGGT CAGAAGACGG TTGCATTTAC	360
CCAGCTACCA TTGCTTCAAT TGATTTTAAG AGAGAAACCT GTGTTGTGGT TTACACTGGA	420
TATGGAAATA GAGAGGAGCA AAATCTGTCC GATCTACTTT CCCCAATCTG TGAAGTAGCT	480
AATAATATAG AACAGAATGC TCAAGAGAAT GAAAATGAAA GCCAAGTTTC AACAGATGAA	540
AGTGAGAACT CCAGGTCTCC TGGAAATAAA TCAGATAACA TCAAGCCCAA ATCTGCTCCA	600
TGGAAACCCT TTCTCCCTCC ACCACCCCCC ATGCCAGGGC CAAGACTGGG ACCAGGAAAG	660
CCAGGTCTAA AATTCAATGG CCCACCACCG CCACCGCCAC CACCACCACC CCACTTACTA	720
TCATGCTGGC TGCCTCCATT TCCTTCTGGA CCACCAATAA TTCCCCCACC ACCTCCCATA	780
TGTCCAGATT CTCTTGATGA TGCTGATGCT TTGGGAAGTA TGTTAATTTT ATGGTACATG	840
AGTGGCTATC ATACTGGCTA TTATATGGGT TTTAGACAAA ATCAAAAAGA AGGAAGGTGC	900
TCACATTCCT TAAATTAAGG AGAAATGCTG GCATAGAGCA GCACTAAATG ACACCACTAA	960
AGAAACGATC AGACAGATCT GGAATGTGAA GCGTTATAGA AGATAACTGG CCTCATTTCT	1020
TCAAAATATC AAGTGTTGGG AAAGAAAAAA GGAAGTGGAA TGGGTAACTC TTCTTGATTA	1080
AAAGTTATGT AATAACCAAA TGCAATGTGA AATATTTTAC TGGACTCTTT TGAAAAACCA	1140
TCTGTAAAAG ACTGAGGTGG GGGTGGGAGG CCAGCACGGT GGTGAGGCAG TTGAGAAAAT	1200
TTGAATGTGG ATTAGATTTT GAATGATATT GGATAATTAT TGGTAATTTT ATGGCCTGTG	1260
AGAAGGGTGT TGTAAGTTTAT AAAAGACTGT CTTAATTTGC ATACTTAAGC ATTTAGGAAT	1320
GAAGTGTTAG AGTGTCTTAA AATGTTTCAA ATGGTTTAAC AAAATGTATG TGAGGCGTAT	1380
GTGGCAAAAT GTTACAGAAT CTAAGTGGTG GACATGGCTG TTCATTGTAC TGTTTTTTTC	1440
TATCTTCTAT ATGTTTAAAA GTATATAATA AAAATATTTA ATTTTTTTTTT AAAAAAAAAA	1500

AAAAAAAAAA AAAAAAAAAA AAAAAAAAAA AAAAAAAAAA AAAAAAAAAA AAAAAAAAAA 1560

AAAAAAAAAA AAAAAAAAAA AA 1582

(2) INFORMATION FOR SEQ ID NO:11:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1408 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:11:

AATTTTAAA	TTTTTGTAG	AGACAGGGTC	TCATTATGTT	GCCCAGGGTG	GTGTCAAGCT	60
CCAGGTCTCA	AGTGATCCCC	CTACCTCCGC	CTCCCAAAGT	TGTGGGATTG	TAGGCATGAG	120
CCACTGCAAG	AAAACCTTAA	CTGCAGCCTA	ATAATTGTTT	TCTTTGGGAT	AACTTTTAAA	180
GTACATTAAA	AGACTATCAA	CTTAATTTCT	GATCATATTT	TGTTGAATAA	AATAAGTAAA	240
ATGTCTTGTG	AACAAAATGC	TTTTTAACAT	CCATATAAAG	CTATCTATAT	ATAGCTATCT	300
ATATCTATAT	AGCTATTTTT	TTTAACTTCC	TTTTATTTTC	CTTACAGGGT	TTTAGACAAA	360
ATCAAAAAGA	AGGAAGGTGC	TCACATTCCT	TAAATTAAGG	AGTAAGTCTG	CCAGCATTAT	420
GAAAGTGAAT	CTTACTTTTG	TAAAACTTTA	TGGTTTGTGG	AAAACAAATG	TTTTTGAACA	480
GTTAAAAAGT	TCAGATGTTA	GAAAGTTGAA	AGGTTAATGT	AAAACAATCA	ATATTAAAGA	540
ATTTTGATGC	CAAAACTATT	AGATAAAAGG	TTAATCTACA	TCCCTACTAG	AATTCTCATA	600
CTTAACTGGT	TGGTTGTGTG	GAAGAAACAT	ACTTTCACAA	TAAAGAGCTT	TAGGATATGA	660
TGCCATTTTA	TATCACTAGT	AGGCAGACCA	GCAGACTTTT	TTTTATTGTG	ATATGGGATA	720
ACCTAGGCAT	ACTGCACTGT	ACACTCTGAC	ATATGAAGTG	CTCTAGTCAA	GTTTAACTGG	780
TGTCCACAGA	GGACATGGTT	TAACTGGAAT	TCGTCAAGCC	TCTGGTTCTA	ATTTCTCATT	840
TGCAGGAAAT	GCTGGCATAG	AGCAGCACTA	AATGACACCA	CTAAAGAAAC	GATCAGACAG	900
ATCTGGAATG	TGAAGCGTTA	TAGAAGATAA	CTGGCCTCAT	TTCTTCAAAA	TATCAAGTGT	960
TGGGAAAGAA	AAAAGGAAGT	GGAATGGGTA	ACTCTTCTTG	ATTAAAAGTT	ATGTAATAAC	1020

CAAATGCAAT GTGAAATATT TTACTGGACT CTTTTGAAAA ACCATCTGTA AAAGACTGAG	1080
GTGGGGGTGG GAGGCCAGCA CGGTGGTGAG GCAGTTGAGA AAATTTGAAT GTGGATTAGA	1140
TTTTGAATGA TATTGGATAA TTATTGGTAA TTTTATGGCC TGTGAGAAGG GTGTTGTAGT	1200
TTATAAAAGA CTGTCTTAAT TTGCATACTT AAGCATTTAG GAATGAAGTG TTAGAGTGTC	1260
TTAAAATGTT TCAAATGGTT TAACAAAATG TATGTGAGGC GTATGTGGCA AAATGTTACA	1320
GAATCTAACT GGTGGACATG GCTGTTCACT GTACTGTTTT TTTCTATCTT CTATATGTTT	1380
AAAAGTATAT AATAAAAATA TTTAATTT	1408

(2) INFORMATION FOR SEQ ID NO:12:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1582 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:12:

CGGGGCCCCA CGCTGCGCAT CCGCGGGTTT GCTATGGCGA TGAGCAGCGG CGGCAGTGGT	60
GGCGGCGTCC CGGAGCAGGA GGATTCCGTG CTGTTCCGGC GCGGCACAGG CCAGAGCGAT	120
GATTCTGACA TTTGGGATGA TACAGCACTG ATAAAAGCAT ATGATAAAGC TGTGGCTTCA	180
TTTAAGCATG CTCTAAAGAA TGGTGACATT TGTGAAACTT CGGGTAAACC AAAAACCACA	240
CCTAAAAGAA AACCTGCTAA GAAGAATAAA AGCCAAAAGA AGAATACTGC AGCTTCCTTA	300
CAACAGTGGA AAGTTGGGGA CAAATGTTCT GCCATTTGGT CAGAAGACGG TTGCATTTAC	360
CCAGCTACCA TTGCTTCAAT TGATTTTAAG AGAGAAACCT GTGTTGTGGT TTACACTGGA	420
TATGGAAATA GAGAGGAGCA AAATCTGTCC GATCTACTTT CCCCAATCTG TGAAGTAGCT	480
AATAATATAG AACAGAATGC TCAAGAGAAT GAAAATGAAA GCCAAGTTTC AACAGATGAA	540
AGTGAGAACT CCAGGTCTCC TGGAAATAAA TCAGATAACA TCAAGCCCAA ATCTGCTCCA	600
TGGAACCTCTT TTCTCCCTCC ACCACCCCCC ATGCCAGGGC CAAGACTGGG ACCAGGAAAG	660
CCAGGTCTAA AATTCAATGG CCCACCACCG CCACCGCCAC CACCACCACC CCACTTACTA	720

TCATGCTGGC TGCCTCCATT TCCTTCTGGA CCACCAATAA TTCCCCCACC ACCTCCCATA	780
TGTCCAGATT CTCTTGATGA TGCTGATGCT TTGGGAAGTA TGTTAATTTT ATGGTACATG	840
AGTGGCTATC ATACTGGCTA TTATATGGGT TTCAGACAAA ATCAAAAAGA AGGAAGGTGC	900
TCACATTCCT TAAATTAAGG AGAAATGCTG GCATAGAGCA GCACTAAATG ACACCACTAA	960
AGAAACGATC AGACAGATCT GGAATGTGAA GCGTTATAGA AGATAACTGG CCTCATTTCT	1020
TCAAAATATC AAGTGTGGG AAAGAAAAA GGAAGTGGAA TGGGTAATC TTCTTGATTA	1080
AAAGTTATGT AATAACCAA TGCAATGTGA AATATTTTAC TGGACTCTTT TGAAAAACCA	1140
TCTGTAAAAG ACTGGGGTGG GGGTGGGAGG CCAGCACGGT GGTGAGGCAG TTGAGAAAAT	1200
TTGAATGTGG ATTAGATTTT GAATGATATT GGATAATTAT TGGTAATTTT ATGGCCTGTG	1260
AGAAAGGGTGT TGTAAGTTTAT AAAAGACTGT CTTAATTTGC ATACTTAAGC ATTTAGGAAT	1320
GAAGTGTTAG AGTGTCTTAA AATGTTTCAA ATGGTTTAAC AAAATGTATG TGAGGCGTAT	1380
GTGGCAAAAT GTTACAGAA CTAACTGGTG GACATGGCTG TTCATTGTAC TGTTTTTTTTC	1440
TATCTTCTAT ATGTTTAAAA GTATATAATA AAAATATTTA ATTTTTTTTT AAAAAAAAAA	1500
AAAAAAAAAA AAAAAAAAAA AAAAAAAAAA AAAAAAAAAA AAAAAAAAAA AAAAAAAAAA	1560
AAAAAAAAAA AAAAAAAAAA AA	1582

(2) INFORMATION FOR SEQ ID NO:13:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1408 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:13:

AATTTTAAAA TTTTGTAG AGACAGGGTC TCATTATGTT GCCCAGGGTG GTGTCAAGCT	60
CCAGGTCTCA AGTGATCCCC CTACCTCCGC CTCCCAAAGT TGTGGGATTG TAGGCATGAG	120
CCACTGCAAG AAAACCTTAA CTGCAGCCTA ATAATTGTTT TCTTTGGGAT AACTTTTAAA	180
GTACATTAAA AGACTATCAA CTTAATTTCT GATCATATTT TGTTGAATAA AATAAGTAAA	240

ATGTCTTGTG AACAAAATGC TTTTAAACAT CCATATAAAG CTATCTATAT ATAGCTATCT	300
ATGTCTATAT AGCTATTTTT TTTAACTTCC TTTTATTTTC CTTACAGGGT TTCAGACAAA	360
ATCAAAAAGA AGGAAGGTGC TCACATTCCT TAAATTAAGG AGTAAGTCTG CCAGCATTAT	420
GAAAGTGAAT CTTACTTTTG TAAAACTTTA TGGTTTGTGG AAAACAAATG TTTTGAACA	480
GTTAAAAAGT TCAGATGTTA AAAAGTTGAA AGGTAAATGT AAAACAATCA ATATTAAAGA	540
ATTTTGATGC CAAAACATTT AGATAAAAGG TTAATCTACA TCCCTACTAG AATTCTCATA	600
CTTAACTGGT TGGTTATGTG GAAGAAACAT ACTTTCACAA TAAAGAGCTT TAGGATATGA	660
TGCCATTTTA TATCACTAGT AGGCAGACCA GCAGACTTTT TTTTATTGTG ATATGGGATA	720
ACCTAGGCAT ACTGCACTGT ACACTCTGAC ATATGAAGTG CTCTAGTCAA GTTTAACTGG	780
TGTCACAGA GGACATGGTT TAACTGGAAT TCGTCAAGCC TCTGGTTCTA ATTTCTCATT	840
TGCAGGAAAT GCTGGCATAG AGCAGCACTA AATGACACCA CTAAAGAAAC GATCAGACAG	900
ATCTGGAATG TGAAGCGTTA TAGAAGATAA CTGGCCTCAT TTCTTCAAAA TATCAAGTGT	960
TGGGAAAGAA AAAAGGAAGT GGAATGGGTA ACTCTTCTTG ATTAAAAGTT ATGTAATAAC	1020
CAAAATGCAAT GTGAAATATT TTAAGTGGT CTTTGAATA ACCATCTGTA AAAGACTGGG	1080
GTGGGGGTGG GAGGCCAGCA CGGTGGTGAG GCAGTTGAGA AAATTTGAAT GTGGATTAGA	1140
TTTGAATGA TATTGGATAA TTATTGGTAA TTTTATGGCC TGTGAGAAGG GTGTTGTAGT	1200
TTATAAAGA CTGTCTTAAT TTGCATACTT AAGCATTTAG GAATGAAGTG TTAGAGTGTC	1260
TTAAATGTT TCAAATGGTT TAACAAAATG TATGTGAGGC GTATGTGGCA AAATGTTACA	1320
GAATCTAACT GGTGGACATG GCTGTTCATT GTACTGTTTT TTTCTATCTT CTATATGTTT	1380
AAAAGTATAT AATAAAAATA TTTAATTT	1408

(2) INFORMATION FOR SEQ ID NO:14:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 360 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

- (A) DESCRIPTION: /desc = "SYNTHETIC DNA"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:14:

ACCTGACCCA GAGGTCAAGG CTGCAGTGAG ACGAGATTGC CCACTGCCCT CCACCCTGGG	60
TGATAAGAGT GGGACCCTGT TCAAAACATA CACACACACA CACACACACA CACACACACA	120
CACACACACA CTCTCTCTCT CTCTCTCTCT CTCTCTCTCT CTCTCTCTCT CTCTCAAAAA	180
CACTTGGTCT GTTATTTTTC GAAATTGTCA GTCATAGTTA TCTGTTAGAC CAAAGCTGGT	240
AAGACATTTA TTACATTGCC TCCTACAAC TCAATCAGCTA ATGTATTTGC TATATAGCAA	300
TTACATATGG ATATATTATC TTAGGGGATG GCCAGTATAA AACTGTCACT GAGGAAAGGA	360

(2) INFORMATION FOR SEQ ID NO:15:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 288 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

- (A) DESCRIPTION: /desc = "SYNTHETIC DNA"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:15:

CTTCCACCT AGCCTCCCCA GTAGCTAGGA CTATAGGCGT GCCCACCAAG CTCAGCTATT	60
TTTTATTTAG TAGAGACGGG GTTTCGGCAG CTTAGGCCTC GTTCGAACTC CAGTGTGTGT	120
GTGTGTGTGT GTGTGTGTGT GTGTGTGTGT GTGTGTGTGT GTGTAGATAT TTATTCCCCC	180
TCCCCCTTGG AAAAGTAAGT AAGCTCCTAC TAGGAATTTA AAACCTGCTT GATCTATATA	240
AAGACAAACA AGGAAAGACA AACATGGGGG CAGGAAGGAA GGCAGATC	288

(2) INFORMATION FOR SEQ ID NO:16:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 141 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

- (A) DESCRIPTION: /desc = "SYNTHETIC DNA"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:16:

TCGAGGTAGA TTTGTATTAT ATCCCATGTA CACACACACA CACACACACA CACACACACA	60
CACACACAGA CTTAATCTGT TTACAGAAAT AAAAGGAATA AAATACCGTT TCTACTATAC	120
ACCAAAACTA GCCATCTTGA C	141

(2) INFORMATION FOR SEQ ID NO:17:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 305 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

- (A) DESCRIPTION: /desc = "SYNTHETIC DNA"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:17:

CCCTGAGAAG GCTTCCTCCT GAGTATGCAT AAACATTAC AGCTTGCATG CGTGTGTGTG	60
TGTGTGTGTG TGTGTATGTT TGCTTGCACT GTAAAAACAA TTGCAACATC AACAGAAATA	120
AAAATTAAAG GAATAATTCT CCTCCGACTC TGCCGTTCCA TCCAGTGAAA CTCTTCATTC	180
TGGGGTAAAG TTCCTTCAGT TCTTTCATAG ATAGGTATAT ACTTCATAAG TCAAACAATC	240
AGGCTGGGTG CAGTAGCTCA TGCCTGTAAT CCCAGCCCTT TGGGAGGCCG AGCTGGGCAG	300
ATCGA	305

(2) INFORMATION FOR SEQ ID NO:18:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 341 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

- (A) DESCRIPTION: /desc = "SYNTHETIC DNA"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:18:

```

TCCACCCGCC TTGGCCTCCC AAAGCCTGGG ATTACAGGCG TGA CTGCCGC ACCCAGCTGT      60
AAACTGGTTT AATGGTAGAT TTTAGGTATT AACAATAGAT AAAAAGATAC TTTTGGCATA      120
CTGTGTATTG GGATGGGGTT AGAACAGGTG TCTACCCAAG ACATT TACTT AAAATCGCCC      180
TCGAAATGCT ATGTGAGCTG TGTGTGTGTG TGTGTGTGTG TGTGTATTAA GGAAAAGCAT      240
GAAAGTATTT ATGCTTGATT TTTTTTTTTC CTCATAGCTT CATAGTGGAC AGATACATAG      300
TCTAAATCAA AATGTTTAAA CTTTTTATGT CACTTGCTGT C                          341

```

(2) INFORMATION FOR SEQ ID NO:19:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 278 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:19:

```

Met Ala Met Ser Ser Gly Gly Ser Gly Gly Gly Val Pro Glu Gln Glu
1          5          10          15
Asp Ser Val Leu Phe Arg Arg Gly Thr Gly Gln Ser Asp Asp Ser Asp
20        25        30
Ile Trp Asp Asp Thr Ala Leu Ile Lys Ala Tyr Asp Lys Ala Val Ala
35        40        45
Ser Phe Lys His Ala Leu Lys Asn Gly Asp Ile Cys Glu Thr Ser Gly
50        55        60
Lys Pro Lys Thr Thr Pro Lys Arg Lys Pro Ala Lys Lys Asn Lys Ser
65        70        75        80
Gln Lys Lys Asn Thr Ala Ala Ser Leu Gln Gln Trp Lys Val Gly Asp
85        90        95
Lys Cys Ser Ala Ile Trp Ser Glu Asp Gly Cys Ile Tyr Pro Ala Thr
100       105       110
Ile Ala Ser Ile Asp Phe Lys Arg Glu Thr Cys Val Val Val Tyr Thr
115       120       125

```

Gly Tyr Gly Asn Arg Glu Glu Gln Asn Leu Ser Asp Leu Leu Ser Pro
 130 135 140
 Ile Cys Glu Val Ala Asn Asn Ile Glu Gln Asn Ala Gln Glu Asn Glu
 145 150 155 160
 Asn Glu Ser Gln Val Ser Thr Asp Glu Ser Glu Asn Ser Arg Ser Pro
 165 170 175
 Gly Asn Lys Ser Asp Asn Ile Lys Pro Lys Ser Ala Pro Trp Asn Ser
 180 185 190
 Phe Leu Pro Pro Pro Pro Pro Met Pro Gly Pro Arg Leu Gly Pro Gly
 195 200 205
 Lys Pro Gly Leu Lys Phe Asn Gly Pro Pro Pro Pro Pro Pro Pro Pro
 210 215 220
 Pro Pro His Leu Leu Ser Cys Trp Leu Pro Pro Phe Pro Ser Gly Pro
 225 230 235 240
 Pro Ile Ile Pro Pro Pro Pro Pro Ile Cys Pro Asp Ser Leu Asp Asp
 245 250 255
 Ala Asp Ala Leu Gly Ser Met Leu Ile Ser Trp Tyr Met Ser Gly Tyr
 260 265 270
 His Thr Gly Tyr Tyr Met
 275

(2) INFORMATION FOR SEQ ID NO:20:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 885 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(ix) FEATURE:

- (A) NAME/KEY: CDS
- (B) LOCATION: 18..881

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:20:

CGGCGTGGTA GCAGGCC ATG GCG ATG GGC AGT GGC GGA GCG GGC TCC GAG
 Met Ala Met Gly Ser Gly Gly Ala Gly Ser Glu
 1 5 10

CAG	GAA	GAT	ACG	GTG	CTG	TTC	CGG	CGT	GGC	ACC	GGC	CAG	AGT	GAT	GAT	98
Gln	Glu	Asp	Thr	Val	Leu	Phe	Arg	Arg	Gly	Thr	Gly	Gln	Ser	Asp	Asp	
			15					20					25			
TCT	GAC	ATT	TGG	GAT	GAT	ACA	GCA	TTG	ATA	AAA	GCT	TAT	GAT	AAA	GCT	146
Ser	Asp	Ile	Trp	Asp	Asp	Thr	Ala	Leu	Ile	Lys	Ala	Tyr	Asp	Lys	Ala	
		30					35					40				
GTG	GCT	TCC	TTT	AAG	CAT	GCT	CTA	AAG	AAC	GGT	GAC	ATT	TGT	GAA	ACT	194
Val	Ala	Ser	Phe	Lys	His	Ala	Leu	Lys	Asn	Gly	Asp	Ile	Cys	Glu	Thr	
	45					50					55					
CCA	GAT	AAG	CCA	AAA	GGC	ACA	GCC	AGA	AGA	AAA	CCT	GCC	AAG	AAG	AAT	242
Pro	Asp	Lys	Pro	Lys	Gly	Thr	Ala	Arg	Arg	Lys	Pro	Ala	Lys	Lys	Asn	
	60				65					70					75	
AAA	AGC	CAA	AAG	AAG	AAT	GCC	ACA	ACT	CCC	TTG	AAA	CAG	TGG	AAA	GTT	290
Lys	Ser	Gln	Lys	Lys	Asn	Ala	Thr	Thr	Pro	Leu	Lys	Gln	Trp	Lys	Val	
				80					85					90		
GGT	GAC	AAG	TGT	TCT	GCT	GTT	TGG	TCA	GAA	GAC	GGC	TGC	ATT	TAC	CCA	338
Gly	Asp	Lys	Cys	Ser	Ala	Val	Trp	Ser	Glu	Asp	Gly	Cys	Ile	Tyr	Pro	
			95				100						105			
GCT	ACT	ATT	ACG	TCC	ATT	GAC	TTT	AAG	AGA	GAA	ACC	TGT	GTC	GTG	GTT	386
Ala	Thr	Ile	Thr	Ser	Ile	Asp	Phe	Lys	Arg	Glu	Thr	Cys	Val	Val	Val	
		110					115					120				
TAT	ACT	GGA	TAT	GGA	AAC	AGA	GAG	GAG	CAA	AAC	TTA	TCT	GAC	CTA	CTT	434
Tyr	Thr	Gly	Tyr	Gly	Asn	Arg	Glu	Glu	Gln	Asn	Leu	Ser	Asp	Leu	Leu	
	125					130					135					
TCC	CCG	ACC	TGT	GAA	GTA	GCT	AAT	AGT	ACA	GAA	CAG	AAC	ACT	CAG	GAG	482
Ser	Pro	Thr	Cys	Glu	Val	Ala	Asn	Ser	Thr	Glu	Gln	Asn	Thr	Gln	Glu	
	140				145					150					155	
AAT	GAA	AGT	CAA	GTT	TCC	ACA	GAC	GAC	AGT	GAA	CAC	TCC	TCC	AGA	TCG	530
Asn	Glu	Ser	Gln	Val	Ser	Thr	Asp	Asp	Ser	Glu	His	Ser	Ser	Arg	Ser	
				160				165						170		
CTC	AGA	AGT	AAA	GCA	CAC	AGC	AAG	TCC	AAA	GCT	GCT	CCG	TGG	ACC	TCA	578
Leu	Arg	Ser	Lys	Ala	His	Ser	Lys	Ser	Lys	Ala	Ala	Pro	Trp	Thr	Ser	
			175					180					185			
TTT	CTT	CCT	CCA	CCA	CCC	CCA	ATG	CCA	GGG	TCA	GGA	TTA	GGA	CCA	GGA	626
Phe	Leu	Pro	Pro	Pro	Pro	Pro	Met	Pro	Gly	Ser	Gly	Leu	Gly	Pro	Gly	
		190					195					200				
AAG	CCA	GGT	CTA	AAA	TTC	AAC	GGC	CCG	CCG	CCG	CCG	CCT	CCA	CTA	CCC	674
Lys	Pro	Gly	Leu	Lys	Phe	Asn	Gly	Pro	Pro	Pro	Pro	Pro	Pro	Pro	Pro	
	205					210					215					

CCT	CCC	CCC	TTC	CTG	CCG	TGC	TGG	ATG	CCC	CCG	TTC	CCT	TCA	GGA	CCA	722
Pro	Pro	Pro	Phe	Leu	Pro	Cys	Trp	Met	Pro	Pro	Phe	Pro	Ser	Gly	Pro	
220					225					230					235	
CCA	ATA	ATC	CCG	CCA	CCC	CCT	CCC	ATC	TCT	CCC	GAC	TGT	CTG	GAT	GAC	770
Pro	Ile	Ile	Pro	Pro	Pro	Pro	Pro	Ile	Ser	Pro	Asp	Cys	Leu	Asp	Asp	
				240					245					250		
ACT	GAT	GCC	CTG	GGC	AGT	ATG	CTA	ATC	TCT	TGG	TAC	ATG	AGT	GGC	TAC	818
Thr	Asp	Ala	Leu	Gly	Ser	Met	Leu	Ile	Ser	Trp	Tyr	Met	Ser	Gly	Tyr	
			255					260					265			
CAC	ACT	GGC	TAC	TAT	ATG	GGT	TTC	AGA	CAA	AAT	AAA	AAA	GAA	GGA	AAG	866
His	Thr	Gly	Tyr	Tyr	Met	Gly	Phe	Arg	Gln	Asn	Lys	Lys	Glu	Gly	Lys	
		270					275					280				
TGC	TCA	CAT	ACA	AAT	TAAG											885
Cys	Ser	His	Thr	Asn												
285																

(2) INFORMATION FOR SEQ ID NO:21:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 288 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:21:

Met	Ala	Met	Gly	Ser	Gly	Gly	Ala	Gly	Ser	Glu	Gln	Glu	Asp	Thr	Val	
1				5					10					15		
Leu	Phe	Arg	Arg	Gly	Thr	Gly	Gln	Ser	Asp	Asp	Ser	Asp	Ile	Trp	Asp	
			20					25					30			
Asp	Thr	Ala	Leu	Ile	Lys	Ala	Tyr	Asp	Lys	Ala	Val	Ala	Ser	Phe	Lys	
		35				40					45					
His	Ala	Leu	Lys	Asn	Gly	Asp	Ile	Cys	Glu	Thr	Pro	Asp	Lys	Pro	Lys	
	50				55						60					
Gly	Thr	Ala	Arg	Arg	Lys	Pro	Ala	Lys	Lys	Asn	Lys	Ser	Gln	Lys	Lys	
65					70					75				80		
Asn	Ala	Thr	Thr	Pro	Leu	Lys	Gln	Trp	Lys	Val	Gly	Asp	Lys	Cys	Ser	
				85				90						95		
Ala	Val	Trp	Ser	Glu	Asp	Gly	Cys	Ile	Tyr	Pro	Ala	Thr	Ile	Thr	Ser	
			100				105						110			

Ile	Asp	Phe	Lys	Arg	Glu	Thr	Cys	Val	Val	Val	Tyr	Thr	Gly	Tyr	Gly
		115					120					125			
Asn	Arg	Glu	Glu	Gln	Asn	Leu	Ser	Asp	Leu	Leu	Ser	Pro	Thr	Cys	Glu
	130					135					140				
Val	Ala	Asn	Ser	Thr	Glu	Gln	Asn	Thr	Gln	Glu	Asn	Glu	Ser	Gln	Val
145					150					155					160
Ser	Thr	Asp	Asp	Ser	Glu	His	Ser	Ser	Arg	Ser	Leu	Arg	Ser	Lys	Ala
				165					170					175	
His	Ser	Lys	Ser	Lys	Ala	Ala	Pro	Trp	Thr	Ser	Phe	Leu	Pro	Pro	Pro
			180					185					190		
Pro	Pro	Met	Pro	Gly	Ser	Gly	Leu	Gly	Pro	Gly	Lys	Pro	Gly	Leu	Lys
		195					200					205			
Phe	Asn	Gly	Pro	Pro	Pro	Pro	Pro	Pro	Leu	Pro	Pro	Pro	Pro	Phe	Leu
	210					215					220				
Pro	Cys	Trp	Met	Pro	Pro	Phe	Pro	Ser	Gly	Pro	Pro	Ile	Ile	Pro	Pro
225				230						235					240
Pro	Pro	Pro	Ile	Ser	Pro	Asp	Cys	Leu	Asp	Asp	Thr	Asp	Ala	Leu	Gly
				245					250					255	
Ser	Met	Leu	Ile	Ser	Trp	Tyr	Met	Ser	Gly	Tyr	His	Thr	Gly	Tyr	Tyr
			260					265					270		
Met	Gly	Phe	Arg	Gln	Asn	Lys	Lys	Glu	Gly	Lys	Cys	Ser	His	Thr	Asn
		275					280					285			

(2) INFORMATION FOR SEQ ID NO:22:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 3246 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:22:

CCTCCCGGGC ACCGTACTGT TCCGCTCCCA GAAGCCCCGG GCGCCGGAAG TCGTCACTCT
 TAAGAAGGGA CGGGGCCCCA CGCTGCGCAC CCGCGGGTTT GCTATGGCGA TGAGCAGCGG

60

120

CGGCAGTGGT	GGCGGCGTCC	CGGAGCAGGA	GGATTCCGTG	CTGTTCCGGC	GCGGCACAGG	180
CAGGTGAGGT	CGCAGCCAGT	GCAGTCTCCC	TATTAGCGCT	CTCAGCACCC	TTCTTCCGGC	240
CCAACTCTCC	TTCCGCAGTA	ATTTTGTAT	GTGTGGATTA	AGATGACTCT	TGGTACTAAC	300
ATACATTTTC	TGATTAAACC	TATCTGACAT	GAGTTGTTTT	TATTTCTTAC	CCTTTCCAGA	360
GCGATGATTC	TGACATTTGG	GATGATACAG	CACTGATAAA	AGCATATGAT	AAAGCTGTGG	420
CTTCATTTAA	GGTATGAAAT	GCTTGTTAGT	CGTTTTCTTA	TTTTCTCGTT	ATTCATTTGG	480
AAAGGAATTG	ATAACATACG	ATAAAGTGTT	AAGTGCTTTC	TGAGGTGACG	GAGCCTTGAG	540
ACTAGCTTAT	AGTAGTAACT	GGGTTATGTC	GTGACTTTTA	TTCTGTGCAC	CACCCTGTAA	600
CATGTACATT	TTTATTCCTA	TTTTCGTAGC	ATGCTCTAAA	GAATGGTGAC	ATTTGTGAAA	660
CTTGGGGTAA	ACCAAAAACC	ACACCTAAAA	GAAAACCTGC	TAAGAAGAAT	AAAAGCCAAA	720
AGAAGAATAC	TGCAGCTTCC	TTACAACAGG	TTATTTTAAA	ATGTTGAGGA	TTTAACTTCA	780
AAGGATGTCT	CATTAGTCCT	TATTTAATAG	TGTAAAATGT	CTTTAACTGC	AGGTCGATCA	840
AAACGAGATG	ATAGTTTGCC	CTCTTCAAAA	GAAATGTGTG	CATGTATATA	TCTTTGATTT	900
CTTTTGTAGT	GGAAAGTTGG	GGACAAATGT	TCTGCCATTT	GGTCAGAAGA	CGGTTGCATT	960
TACGCAGCTA	CCATTGCTTC	AATTGATTTT	AAGAGAGAAA	CCTGTGTTGT	GGTTTACACT	1020
GGATATGGAA	ATAGAGAGGA	GCAAAATCTG	TCCGATCTAC	TTTCCCCAAT	CTGTGAAGTA	1080
GCTATAATA	TAGAACAGAA	TGCTCAAGAG	GTAAGGATAC	AAAAAAAAAA	AAATTCAATT	1140
TCTGGAAGCA	GAGACTAGAT	GAGAAACTGT	TAAACAGTAT	ACAACCGAGG	CATTAATTTT	1200
TTCTTAATCA	CACCCTTATA	ACAAAAACCT	GCATATTTTT	TCTTTTAAA	GAATGAAAAT	1260
GAAAGCCAAG	TTTCAACAGA	TGAAAGTGAG	AACTCCAGGT	CTCCTGGAAA	TAAATCAGAT	1320
AACATCAAGC	CCAAATCTGC	TCCATGGAAC	TCTTTTCTCC	CTCCACCACC	CCCCATGCCA	1380
GGGCCAAGAC	TGGGACCAGG	AAAGGTAAAC	CTTCTATGAA	AGTTTTCCAG	AAAATAGTTA	1440
ATGTCGGGAC	ATTTAACCTC	TCTGTTAACT	AATTTGTAGC	TCTCCAATAT	TCTGGGTAAT	1500
TATTTTTATC	CTTTTGGTTT	TGAGTCCTTT	TTATTCCTAT	CATATTGAAA	TTGGTAAGTT	1560
AATTTTCCTT	TGAAATATTC	CTTATAGCCA	GGTCTAAAAT	TCAATGGCCC	ACCACCGCCA	1620
CCGCCACCAC	CACCACCCCA	CTTACTATCA	TGCTGGCTGC	CTCCATTTCC	TTCTGGACCA	1680

CCAGTAAGTA	AAAAAGAGTA	TAGGTTAGAT	TTTGCTTTCA	CATACAATTT	GATAATAGAC	1740
TTTACTTTTT	GTTTACTGGA	TATAACAAT	ATCTTTTCT	GTCTCCAGAT	AATTCCCCCA	1800
CCACCTCCCA	TATGTCCAGA	TTCTCTTGAT	GATGCTGATG	CTTTGGGAAG	TATGTTAATT	1860
TCATGGTACA	TGAGTGGCTA	TCATACTGGC	TATTATATGG	TAAGTAATCA	CTCAGCATCT	1920
TTTCCTGACA	ATTTTTTTGT	AGTTATGTGA	CTTTGTTTGG	TAAATTTATA	AAATACTACT	1980
CTGCAGCCTA	ATAATTGTTT	TCTTTGGGAT	AACTTTTAAA	GTACATTAAA	AGACTATCAA	2040
CTTAATTTCT	GATCATATTT	TGTTGAATAA	AATAAGTAAA	ATGTCTTGTG	AAACAAAATG	2100
CTTTTAAACA	TCCATATAAA	GCTATCTATA	TATAGCTATC	TATGTCTATA	TAGCTATTTT	2160
TTTTAACTTC	CTTTTATTTT	CCTTACAGGG	TTTCAGACAA	AATCAAAAAG	AAGGAAGGTG	2220
CTCACATTCC	TTAAATTAAG	GAGTAAGTCT	GCCAGCATT	TGAAAGTGAA	TCTTACTTTT	2280
GTAAAACTTT	ATGGTTTGTG	GAAAACAAAT	GTTTTTGAAC	AGTTAAAAAG	TTCAGATGTT	2340
AAAAAGTTGA	AAGGTTAATG	TAAAACAATC	AATATTAAAG	AATTTTGATG	CCAAAACAT	2400
TAGATAAAAG	GTTAATCTAC	ATCCCTACTA	GAATTCTCAT	ACTTAACTGG	TTGGTTATGT	2460
GGAAGAAACA	TACTTTCACA	ATAAAGAGCT	TTAGGATATG	ATGCCATTTT	ATATCACTAG	2520
TAGGCAGACC	AGCAGACTTT	TTTTTATTGT	GATATGGGAT	AACCTAGGCA	TACTGCACTG	2580
TAGACTCTGA	CATATGAAGT	GCTCTAGTCA	AGTTTAACTG	GTGTCCACAG	AGGACATGGT	2640
TTAACTGGAA	TTCGTCAAGC	CTCTGGTTCT	AATTTCTCAT	TTGCAGGAAA	TGCTGGCATA	2700
GAGCAGCACT	AAATGACACC	ACTAAAGAAA	CGATCAGACA	GATCTGGAAT	GTGAAGCGTT	2760
ATAGAAGATA	ACTGGCCTCA	TTTCTTCAAA	ATATCAAGTG	TTGGGAAAGA	AAAAAGGAAG	2820
TGGAATGGGT	AACTCTTCTT	GATTAAAAGT	TATGTAATAA	CCAAATGCAA	TGTGAAATAT	2880
TTTACTGGAC	TCTTTTGAAA	AACCATCTAG	TAAAAGACTG	GGGTGGGGGT	GGGAGGCCAG	2940
CACGGTGGTG	AGGCAGTTGA	GAAAATTTGA	ATGTGGATTA	GATTTTGAAT	GATATTGGAT	3000
AATTATTGGT	AATTTTATGG	CCTGTGAGAA	GGGTGTTGTA	GTTTATAAAA	GACTGTCTTA	3060
ATTTGCATAC	TTAAGCATTT	AGGAATGAAG	TGTTAGAGTG	TCTTAAAATG	TTTCAAATGG	3120
TTTAACAAAA	TGTATGTGAG	GCGTATGTGG	CAAATGTTA	CAGAATCTAA	CTGGTGGACA	3180
TGGCTGTTCA	TTGTACTGTT	TTTTTCTATC	TTCTATATGT	TTAAAAGTAT	ATAATAAAAA	3240

TATTTA

3246

(2) INFORMATION FOR SEQ ID NO:23:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 637 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:23:

GATCTGCCTT CCTTCCTGCC CCCATGTTTG TCTTTCCTTG TTTGTCTTTA TATAGATCAA	60
GCAGGTTTTA AATTCCTAGT AGGAGCTTAC ATTTACTTTT CCAAGGGGGA GGGGGAATAA	120
ATATCTACAC ACACACACAC ACACACACCA CACTGGAGTT CGAGACGAGG CCTAAGCAAC	180
ATGCCGAAAC CCCGTCTCTA CTAAATACAA AAAATAGCTG AGCTTGGTGG CGCACGCCTA	240
TAGTCCTAGC TACTGGGGAG GCTGAGGTGG GAGGATCGCT TGAGCCCAAG AAGTCGAGGC	300
TGCAGTGAGC CGAGATCGCG CCGCTGCACT CCAGCCTGAG CGACAGGGCG AGGCTCTGTC	360
TCAAACAAA CAAACAAAAA AAAAAAGGAA AGGAAATATA ACACAGTGAA ATGAAAGGAT	420
TGAGAGAAAT GAAAAATATA CACGCCACAA ATGTGGGAGG GCGATAACCA CTCGTAGAAA	480
GCGTGAGAAG TTAACAAG CGGTCCTCCC GGGCACCGTA CTGTTCCGCT CCCAGAAGCC	540
CCGGGCGCCG GAAGTCGTCA CTCTTAAGAA GGGACGGGGC CCCACGCTGC GCACCCGCGG	600
GTTTGCTATG GCGATGAGCA GCGGCGGCAG TGGTGGC	637

(2) INFORMATION FOR SEQ ID NO:24:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 19 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

(iii) HYPOTHETICAL: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:24:

AGGGCGAGGC TCTGTCTCA

19

(2) INFORMATION FOR SEQ ID NO:25:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 19 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

(iii) HYPOTHETICAL: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:25:

CGGGAGGACC GCTTGTAGT

19

(2) INFORMATION FOR SEQ ID NO:26:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 19 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

(iii) HYPOTHETICAL: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:26:

GCCGGAAGTC GTCACCTCTT

19

(2) INFORMATION FOR SEQ ID NO:27:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 20 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

(iii) HYPOTHETICAL: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:27:

GGGTGCTGAG AGCGCTAATA

20

(2) INFORMATION FOR SEQ ID NO:28:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 20 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

(iii) HYPOTHETICAL: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:28:

TGTTGGATT AAGATGACTC

20

(2) INFORMATION FOR SEQ ID NO:29:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 19 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

(iii) HYPOTHETICAL: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:29:

CACTTTATCG TATGTTATC

19

(2) INFORMATION FOR SEQ ID NO:30:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 22 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

(iii) HYPOTHETICAL: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:30:

CTGTGCACCA CCCTGTAACA TG

22

(2) INFORMATION FOR SEQ ID NO:31:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 19 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: other nucleic acid
- (iii) HYPOTHETICAL: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:31:

AAGGACTAAT GAGACATCC

19

(2) INFORMATION FOR SEQ ID NO:32:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 20 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: other nucleic acid
- (iii) HYPOTHETICAL: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:32:

CGAGATGATA GTTTGCCCTC

20

(2) INFORMATION FOR SEQ ID NO:33:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 24 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

(iii) HYPOTHETICAL: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:33:

AGCTACTTCA CAGATTGGGG AAAG

24

(2) INFORMATION FOR SEQ ID NO:34:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 20 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

(iii) HYPOTHETICAL: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:34:

CTCATCTAGT CTCTGCTTCC

20

(2) INFORMATION FOR SEQ ID NO:35:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 24 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

(iii) HYPOTHETICAL: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:35:

TGGATATGGA AATAGAGAGG GAGC

24

(2) INFORMATION FOR SEQ ID NO:36:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 22 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

(iii) HYPOTHETICAL: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:36:

CACCCTTATA ACAAAAACCT GC

22

(2) INFORMATION FOR SEQ ID NO:37:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 23 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

(iii) HYPOTHETICAL: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:37:

GAGAAAGGAG TTCCATGGAG CAG

23

(2) INFORMATION FOR SEQ ID NO:38:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 20 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

(iii) HYPOTHETICAL: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:38:

GAGAGGTTAA ATGTCCCGAC

20

(2) INFORMATION FOR SEQ ID NO:39:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 22 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

(iii) HYPOTHETICAL: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:39:

GTGAGAACTC CAGGTCTCCT GG

22

(2) INFORMATION FOR SEQ ID NO:40:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 20 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

(iii) HYPOTHETICAL: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:40:

TGAGTCTGTT TGAATTCAGG

20

(2) INFORMATION FOR SEQ ID NO:41:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 22 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

(iii) HYPOTHETICAL: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:41:

GAAGGAAATG GAGGCAGCCA GC

22

(2) INFORMATION FOR SEQ ID NO:42:

- SEQUENCE
- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 21 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear
 - (ii) MOLECULE TYPE: other nucleic acid
 - (iii) HYPOTHETICAL: NO

SEQUENCE

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:42:

TTTCTACCCA TTAGAATCTG G

21

(2) INFORMATION FOR SEQ ID NO:43:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 24 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: other nucleic acid
- (iii) HYPOTHETICAL: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:43:

CCCCACTTAC TATCATGCTG GCTG

24

(2) INFORMATION FOR SEQ ID NO:44:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 24 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

(iii) HYPOTHETICAL: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:44:

CCAGACTTTA CTTTTTGTTT ACTG

24

(2) INFORMATION FOR SEQ ID NO:45:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 21 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

(iii) HYPOTHETICAL: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:45:

ATAGCCACTC ATGTACCATG A

21

(2) INFORMATION FOR SEQ ID NO:46:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 24 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

(iii) HYPOTHETICAL: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:46:

AAGAGTAATT TAAGCCTCAG ACAG

24

(2) INFORMATION FOR SEQ ID NO:47:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 24 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

(iii) HYPOTHETICAL: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:47:

CTGCCATATG TCCAGATTCT CTTG

24

(2) INFORMATION FOR SEQ ID NO:48:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 25 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

(iii) HYPOTHETICAL: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:48:

AGACTATCAA CTTAATTTCT GATCA

25

(2) INFORMATION FOR SEQ ID NO:49:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 24 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

(iii) HYPOTHETICAL: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:49:

TAAGGAATGT GAGCACCTTC CTTC

24

(2) INFORMATION FOR SEQ ID NO:50:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 25 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

(iii) HYPOTHETICAL: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:50:

AGACTATCAA CTTAATTCT GATCA

25

(2) INFORMATION FOR SEQ ID NO:51:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 24 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

(iii) HYPOTHETICAL: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:51:

GTAAGATTCA CTTTCATAAT GCTG

24

(2) INFORMATION FOR SEQ ID NO:52:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 21 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

(iii) HYPOTHETICAL: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:52:

CTTTATGGTT TGTGGAAAAC A

21

(2) INFORMATION FOR SEQ ID NO:53:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 20 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

(iii) HYPOTHETICAL: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:53:

GGCATCATAT CCTAAAGCTC

20

(2) INFORMATION FOR SEQ ID NO:54:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 23 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

(iii) HYPOTHETICAL: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:54:

GTAATAACCA AATGCAATGT GAA

23

(2) INFORMATION FOR SEQ ID NO:55:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 20 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

(iii) HYPOTHETICAL: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:55:

CTACAACACC CTTCTCACAG

20

(2) INFORMATION FOR SEQ ID NO:56:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 20 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

(iii) HYPOTHETICAL: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:56:

GGTGTCCACA GAGGACATGG

20

(2) INFORMATION FOR SEQ ID NO:57:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 24 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

(iii) HYPOTHETICAL: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:57:

AAGAGTTAAC CCATTCCAGC TTCC

24

1 AAGAGTTAAC CCATTCCAGC TTCC
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841
842
843
844
845
846
847
848
849
850
851
852
853
854
855
856
857
858
859
860
861
862
863
864
865
866
867
868
869
870
871
872
873
874
875
876
877
878
879
880
881
882
883
884
885
886
887
888
889
890
891
892
893
894
895
896
897
898
899
900
901
902
903
904
905
906
907
908
909
910
911
912
913
914
915
916
917
918
919
920
921
922
923
924
925
926
927
928
929
930
931
932
933
934
935
936
937
938
939
940
941
942
943
944
945
946
947
948
949
950
951
952
953
954
955
956
957
958
959
960
961
962
963
964
965
966
967
968
969
970
971
972
973
974
975
976
977
978
979
980
981
982
983
984
985
986
987
988
989
990
991
992
993
994
995
996
997
998
999
1000
1001
1002
1003
1004
1005
1006
1007
1008
1009
1010
1011
1012
1013
1014
1015
1016
1017
1018
1019
1020
1021
1022
1023
1024
1025
1026
1027
1028
1029
1030
1031
1032
1033
1034
1035
1036
1037
1038
1039
1040
1041
1042
1043
1044
1045
1046
1047
1048
1049
1050
1051
1052
1053
1054
1055
1056
1057
1058
1059
1060
1061
1062
1063
1064
1065
1066
1067
1068
1069
1070
1071
1072
1073
1074
1075
1076
1077
1078
1079
1080
1081
1082
1083
1084
1085
1086
1087
1088
1089
1090
1091
1092
1093
1094
1095
1096
1097
1098
1099
1100
1101
1102
1103
1104
1105
1106
1107
1108
1109
1110
1111
1112
1113
1114
1115
1116
1117
1118
1119
1120
1121
1122
1123
1124
1125
1126
1127
1128
1129
1130
1131
1132
1133
1134
1135
1136
1137
1138
1139
1140
1141
1142
1143
1144
1145
1146
1147
1148
1149
1150
1151
1152
1153
1154
1155
1156
1157
1158
1159
1160
1161
1162
1163
1164
1165
1166
1167
1168
1169
1170
1171
1172
1173
1174
1175
1176
1177
1178
1179
1180
1181
1182
1183
1184
1185
1186
1187
1188
1189
1190
1191
1192
1193
1194
1195
1196
1197
1198
1199
1200
1201
1202
1203
1204
1205
1206
1207
1208
1209
1210
1211
1212
1213
1214
1215
1216
1217
1218
1219
1220
1221
1222
1223
1224
1225
1226
1227
1228
1229
1230
1231
1232
1233
1234
1235
1236
1237
1238
1239
1240
1241
1242
1243
1244
1245
1246
1247
1248
1249
1250
1251
1252
1253
1254
1255
1256
1257
1258
1259
1260
1261
1262
1263
1264
1265
1266
1267
1268
1269
1270
1271
1272
1273
1274
1275
1276
1277
1278
1279
1280
1281
1282
1283
1284
1285
1286
1287
1288
1289
1290
1291
1292
1293
1294
1295
1296
1297
1298
1299
1300
1301
1302
1303
1304
1305
1306
1307
1308
1309
1310
1311
1312
1313
1314
1315
1316
1317
1318
1319
1320
1321
1322
1323
1324
1325
1326
1327
1328
1329
1330
1331
1332
1333
1334
1335
1336
1337
1338
1339
1340
1341
1342
1343
1344
1345
1346
1347
1348
1349
1350
1351
1352
1353
1354
1355
1356
1357
1358
1359
1360
1361
1362
1363
1364
1365
1366
1367
1368
1369
1370
1371
1372
1373
1374
1375
1376
1377
1378
1379
1380
1381
1382
1383
1384
1385
1386
1387
1388
1389
1390
1391
1392
1393
1394
1395
1396
1397
1398
1399
1400
1401
1402
1403
1404
1405
1406
1407
1408
1409
1410
1411
1412
1413
1414
1415
1416
1417
1418
1419
1420
1421
1422
1423
1424
1425
1426
1427
1428
1429
1430
1431
1432
1433
1434
1435
1436
1437
1438
1439
1440
1441
1442
1443
1444
1445
1446
1447
1448
1449
1450
1451
1452
1453
1454
1455
1456
1457
1458
1459
1460
1461
1462
1463
1464
1465
1466
1467
1468
1469
1470
1471
1472
1473
1474
1475
1476
1477
1478
1479
1480
1481
1482
1483
1484
1485
1486
1487
1488
1489
1490
1491
1492
1493
1494
1495
1496
1497
1498
1499
1500
1501
1502
1503
1504
1505
1506
1507
1508
1509
1510
1511
1512
1513
1514
1515
1516
1517
1518
1519
1520
1521
1522
1523
1524
1525
1526
1527
1528
1529
1530
1531
1532
1533
1534
1535
1536
1537
1538
1539
1540
1541
1542
1543
1544
1545
1546
1547
1548
1549
1550
1551
1552
1553
1554
1555
1556
1557
1558
1559
1560
1561
1562
1563
1564
1565
1566
1567
1568
1569
1570
1571
1572
1573
1574
1575
1576
1577
1578
1579
1580
1581
1582
1583
1584
1585
1586
1587
1588
1589
1590
1591
1592
1593
1594
1595
1596
1597
1598
1599
1600
1601
1602
1603
1604
1605
1606
1607
1608
1609
1610
1611
1612
1613
1614
1615
1616
1617
1618
1619
1620
1621
1622
1623
1624
1625
1626
1627
1628
1629
1630
1631
1632
1633
1634
1635
1636
1637
1638
1639
1640
1641
1642
1643
1644
1645
1646
1647
1648
1649
1650
1651
1652
1653
1654
1655
1656
1657
1658
1659
1660
1661
1662
1663
1664
1665
1666
1667
1668
1669
1670
1671
1672
1673
1674
1675
1676
1677
1678
1679
1680
1681
1682
1683
1684
1685
1686
1687
1688
1689
1690
1691
1692
1693
1694
1695
1696
1697
1698
1699
1700
1701
1702
1703
1704
1705
1706
1707
1708
1709
1710
1711
1712
1713
1714
1715
1716
1717
1718
1719
1720
1721
1722
1723
1724
1725
1726
1727
1728
1729
1730
1731
1732
1733
1734
1735
1736
1737
1738
1739
1740
1741
1742
1743
1744
1745
1746
1747
1748
1749
1750
1751
1752
1753
1754
1755
1756
1757
1758
1759
1760
1761
1762
1763
1764
1765
1766
1767
1768
1769
1770
1771
1772
1773
1774
1775
1776
1777
1778
1779
1780
1781
1782
1783
1784
1785
1786
1787
1788
1789
1790
1791
1792
1793
1794
1795
1796
1797
1798
1799
1800
1801
1802
1803
1804
1805
1806
1807
1808
1809
1810
1811
1812
1813
1814
1815
1816
1817
1818
1819
1820
1821
1822
1823
1824
1825
1826
1827
1828
1829
1830
1831
1832
1833
1834
1835
1836
1837
1838
1839
1840
1841
1842
1843
1844
1845
1846
1847
1848
1849
1850
1851
1852
1853
1854
1855
1856
1857
1858
1859
1860
1861
1862
1863
1864
1865
1866
1867
1868
1869
1870
1871
1872
1873
1874
1875
1876
1877
1878
1879
1880
1881
1882
1883
1884
1885
1886
1887
1888
1889
1890
1891
1892
1893
1894
1895
1896
1897
1898
1899
1900
1901
1902
1903
1904
1905
1906
1907
1908
1909
1910
1911
1912
1913
1914
1915
1916
1917
1918
1919
1920
1921
1922
1923
1924
1925
1926
1927
1928
1929
1930
1931
1932
1933
1934
1935
1936
1937
1938
1939
1940
1941
1942
1943
1944
1945
1946
1947
1948
1949
1950
1951
1952
1953
1954
1955
1956
1957
1958
1959
1960
1961
1962
1963
1964
1965
1966
1967
1968
1969
1970
1971
1972
1973
1974
1975
1976
1977
1978
1979
1980
1981
1982
1983
1984
1985
1986
1987
1988
1989
1990
1991
1992
1993
1994
1995
1996
1997
1998
1999
2000
2001
2002
2003
2004
2005
2006
2007
2008
2009
2010
2011
2012
2013
2014
2015
2016
2017
2018
2019
2020
2021
2022
2023
2024
2025
2026
2027
2028
2029
2030
2031
2032
2033
2034
2035
2036
2037
2038
2039
2040
2041
2042
2043
2044
2045
2046
2047
2048
2049
2050
2051
2052
2053
2054
2055
2056
2057
2058
2059
2060
2061
2062
2063
2064
2065
2066
2067
2068
2069
2070
2071
2072
2073
2074
2075
2076
2077
2078
2079
2080
2081
2082
2083
2084
2085
2086
2087
2088
2089
2090
2091
2092
2093
2094
2095
2096
2097
2098
2099
2100
2101
2102
2103
2104
2105
2106
2107
2108
2109
2110
2111
2112
2113
2114
2115
2116
2117
2118
2119
2120
2121
2122
2123
2124
2125
2126
2127
2128
2129
2130
2131
2132
2133
2134
2135
2136
2137
2138
2139
2140
2141
2142
2143
2144
2145
2146
2147
2148
2149
2150
2151
2152
2153
2154
2155
2156
2157
2158
2159
2160
2161
2162
2163
2164
2165
2166
2167
2168
2169
2170
2171
2172
2173
2174
2175
2176
2177
2178
2179
2180
2181
2182
2183
2184
2185
2186
2187
2188
2189
2190
2191
2192
2193
2194
2195
2196
2197
2198
2199
2200
2201
2202
2203
2204
2205
2206
2207
2208
2209
2210
2211
2212
2213
2214
2215
2216
2217